yw

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/066,463	01/31/2002	Hideaki Kurihara	FUJO 19.398	2188

7590 02/10/2005

Rosenman & Colin LLP
575 Madison Avenue
New York, NY 10022-2585

EXAMINER	
LERNER, MARTIN	

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 02/10/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/066,463	KURIHARA ET AL.	
	Examiner	Art Unit	
	Martin Lerner	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *13 December 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *2 to 4, 6 to 8, and 10 to 15* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *2 to 4, 6 to 8, and 10 to 15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *09/24/2004*.
4) ☒ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 2, 3, 6, 7, 10, and 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Lee et al. ("A Study on a Reduction of the Transmission Bit Rate by U/V Decision Using LSP in the CELP Vocoder")* in view of *Gersho et al. ("An Overview of Variable Rate Speech Coding for Cellular Networks")*.

Regarding independent claims 2, 6, and 10, *Lee et al.* discloses a CELP vocoder device, method, and computer program, comprising:

"an LSP coefficient calculating unit calculating an LSP coefficient obtained from the voice signal" – line spectral pairs (LSPs) are calculated by LPC analysis of speech signal S (Pages 997, Right Column to Page 998, Right Column, II: Calculation of the LSP); Figure 4 shows a flowchart of the process includes a step called Extraction of LSP parameters (Page 999, Right Column: Figure 4);

"an LSP interval judging unit judging whether an interval on a frequency axis between the LSP coefficients is equal to or less than a prescribed threshold" – *int v(i)* is the LSP interval, where *int v(i)* = $| p_{i+1} - p_i |$, for a vector of LSPs $P = [p_1, p_2, \ldots, p_{10}]$

(Page 999, Right Column: Equation (18)); a test is made to determine whether min *int v(i)*, the minimum interval between line spectral pairs in an LSP interval vector *int v(i)*, is less than $F_s/4$, where $F_s/4$ is the threshold ("equal to or less than a prescribed threshold"); Figure 4 shows a flowchart of the process includes a step determining whether min *int v(i)* < $F_s/4$ (Page 999, Right Column: Figure 4); LSPs [p_1, p_2, . . ., p_{10}] are points on a frequency axis, so *int v(i)* are intervals on a frequency axis;

"a judging unit judging whether a voice signal is a vowel when a voice part of a voice signal is sounded" – to decide U/V, the NL and the NH value are detected; in the case of NL is larger than NH, the speech spectrum is represented as a voiced speech spectrum; thus, the frame is decided to be voiced speech; in the other case of NH is larger than NL, the frame is decided to be unvoiced speech; that is, the unvoiced speech has formants in a high frequency band; however, some vowels' NH is larger than NL because vowels such as /i/, /I/, /ɛ/, /æ/ have high second, third, and fourth formants; such frames are decided by the existence of the first formant; if the LSP intervals are detected and are narrow, the frame is decided to be voiced sounds (Page 999, Right Column: Figure 4); thus, LSP intervals are employed to make special arrangements for some vowels by considering whether *int v(i)* < $F_s/4$ and a < b so the frame can be correctly classified as voiced; Figure 4 shows a flowchart of the process includes steps determining whether *int v(i)* < $F_s/4$ and a < b for these vowels (Page 999, Right Column: Figure 4).

Lee et al. discloses reduction of a transmission bit rate by U/V decision using LSP parameters when testing for some vowels. (Page 1000: Table 1) An overall bit

rate can be reduced because unvoiced portions can be encoded with 32 bits. However, *Lee et al.* does not specifically disclose a rate setting unit setting a voice encoding bit rate to a lower bit rate when a vowel is present. That is, *Lee et al.* omits "a rate setting unit setting a voice encoding bit rate, if the voice signal is a vowel said voice encoding bit rate is set to a bit rate lower than the bit rate usually used when the voice part is sounded." Still, variable rate speech coding is fairly well known for reducing an overall bit rate by encoding voiced and unvoiced sounds with different encoding algorithms. *Gersho et al.* teaches voice activity controlled variable rate coding, and particularly a Phonetically Segmented VXC, where each coding frame is analyzed to determine a set of features that are then used to phonetically classify the frame. A variable coding rate is set for different phonetic segments. Bits can also be saved in encoding sustained vowels sounds. (Page 174, Left Column) Thus, *Gersho et al.* suggests variable rate speech coding for phonetic segments including certain vowels in order to reduce the overall bit rate. It would have been obvious to one having ordinary skill in the art to include a rate setting unit setting a voice encoding bit rate to a lower bit rate when certain vowels are detected as taught by *Gersho et al.* in the LSP CELP vocoder of *Lee et al.* for the purpose of reducing the overall bit rate by changing the encoding algorithm for certain vowels.

Regarding claims 3, 7, and 11, *Gersho et al.* teaches a variable coding rate is set for different phonetic segments, where bits can also be saved in encoding sustained vowels sounds (Page 174, Left Column); a "sustained vowel" presumes parameters of

the speech signal (i.e. LSPs) for the vowel do not move and are constant for a given time period; also, *Gersho et al.* teaches switching between various rates based on whether a short-term quality measure remains constant as a function of time (Page 174, Left Column, First Full Paragraph, citing Lundheim and Ramstad).

3. Claims 4, 8, and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Lee et al.* in view of *Gersho et al.* as applied to claims 2, 6, and 10 above, and further in view of *Kang et al.*

Neither *Lee et al.* nor *Gersho et al.* disclose using templates to determine whether a speech segment is a vowel, although templates are well known for identifying the phonetic content of a speech segment by comparing speech segment parameters to parameters representing a class of phonetic features stored in the template. *Kang et al.* teaches a voice communication processing system, where a filter coefficient table contains line spectrum pair (LSP) sets, and particularly filter coefficient templates representing vowels by line spectral frequencies. It is suggested representing speech parameters by LSP-based templates has the advantage of reducing the bit rate. (Column 5, Line 67 to Column 7, Line 50, and particularly Column 6, Line 66 to Column 7, Line 50) It would have been obvious to one having ordinary skill in the art to determine whether a speech segment is a vowel by comparing to templates of LSP coefficients as taught by *Kang et al.* in the LSP CELP vocoder of *Lee et al.* for the purpose of reducing a bit rate.

4. Claims 13 to 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Lee et al.* in view of *Gersho et al.* as applied to claims 2, 6, and 10 above, and further in view of *Das*.

Gersho et al. suggests that most of the coders in the TIA half-rate assessment have incorporated some type of phonetic segmentation. (Page 174, Last Full ¶) However, *Gersho et al.* omits specific disclosure of setting the encoding bit rate at half the usual bit rate when the voice part is a vowel. *Das* teaches a multimode speech coder, where voiced speech frames with sufficient periodicity are encoded spectrally at half rate, or 4 kbps. (Column 8, Lines 35 to 53: Figure 5: Steps 408, 412, and 416) The multimode coder makes a decision as to whether the frame is transition (T), voiced (V), unvoiced (U), or noise (N). If the frame is voiced (V), then the speech is processed under V mode, i.e. at half rate. The stated advantage is that the high-bit-rate T mode is used only when necessary, exploiting the periodicity of voiced speech segments with the lower-bit-rate V mode while preventing any lapse in quality by switching to full rate when the V mode does not perform adequately. (Column 14, Line 21 to Column 15, Line 9: Figure 9) Those skilled in the art would know that a vowel is the most common example of purely voiced speech, and has the most periodicity. Thus, *Das* suggests the bit rate can be reduced to half rate when the frame is voiced, which is commonly a vowel. It would have been obvious to one having ordinary skill in the art to set the voice encoding bit rate to half rate for a voiced frame, which is a vowel, as suggested by *Das* in the CELP vocoder of *Lee et al.* for the purpose of reducing the bit rate in a multimode coder without sacrificing voice quality.

Response to Arguments

5. Applicants' arguments filed 13 December 2004 have been fully considered but they are not persuasive.

Firstly, Applicants point out that the Official Action admits that *Lee et al.* does not disclose "a rate setting unit setting a voice encoding bit rate, if the voice signal is a vowel said voice encoding bit rate is set to a bit rate lower than the bit rate usually used when the voice part is sounded."

However, while *Lee et al.* does not expressly disclose a rate setting unit, in fact, the reference does implicitly suggest the possibility of a rate setting unit to one skilled in the art. *Lee et al.* discloses a U/V decision algorithm performed in a 5.3 kbps ACELP vocoder, where the frame size is 240 msec and the subframe size is 60 msec. (Page 999: Figure 3) Each frame is classified as voiced or unvoiced by the proposed decision algorithm, and an unvoiced frame is encoded with a total of 32 bits. (Page 999: Figure 4) An overall reduction of transmission bit rate is approximately 10%. (Page 1000: Table 1) Thus, *Lee et al.*'s U/V decision algorithm operates on the basis of segments consisting of frames, where each voiced frame is encoded with fewer bits than for an unvoiced frame. Encoding voiced frames with fewer bits as compared to encoding unvoiced frames with a greater number of bits is equivalent to changing a bit rate with respect to frame-sized units. One skilled in the art would know that changing a number of bits for encoding frame segments is equivalent to "a rate setting unit setting a voice

encoding bit rate . . . to a bit rate lower than the bit rate usually used when the voice part is sounded."

Secondly, Applicants argue that *Lee et al.* does not teach judging a vowel is present in a *voice signal* when an interval on a frequency axis between coefficients is equal to or less than a prescribed threshold. Applicants maintain that *Lee et al.* only states that *certain* vowel sounds exhibit characteristics of an un-voiced part, but does not teach judging a vowel is present *in the voice signal*. (*Applicants' emphasis*) This position is traversed.

Applicants' use of the term "voice signal" is ambiguous and inconsistent. The language of the claims does not make it expressly clear whether a "voice signal" refers to a speech signal or a voiced signal. Generally, a speech signal may be classified (at least) into segments that are silence, voiced, and unvoiced. The term "voice signal" may refer to a speech signal that is not silence, or may refer to an entirety of a signal containing speech in both silent and non-silent (speech) portions. Or, the term "voice signal" may refer to only a voiced portion of non-silent speech, as distinguished from unvoiced portions. (Typically, those skilled in the art know that voiced portions correspond to vowel sounds, and unvoiced portions correspond to consonant sounds.) It is unclear whether Applicants intend for the term "voice signal" to refer only to voiced sounds. However, the claims do not say "voiced", only "voice", which is a different word. During patent examination, the pending claims must be "given their broadest reasonable interpretation consistent with the specification." *In re Hyatt*, 211 F.3d 1367, 1372, 54 USPQ2d 1664, 1667 (Fed. Cir. 2000). Applicant always has the opportunity to

amend the claims during prosecution, and broad interpretation by the examiner reduces the possibility that the claim, once issued, will be interpreted more broadly than is justified. *In re Prater*, 415 F.2d 1393, 1404-05, 162 USPQ 541, 550-51 (CCPA 1969) Here, it is unclear whether Applicants intend the term "voice signal" to refer to a speech signal, generally, or more specifically, to a voiced signal. Under principles of broadest reasonable interpretation, though, one skilled in the art would read "voice signal" as meaning "speech signal".

Lee et al. clearly discloses judging whether segments of a speech signal are a vowel by comparing an interval on a frequency axis between LSP coefficients as being equal to or less than a prescribed threshold. LSP's are line spectral pairs on a frequency axis. It is stated "some vowels' NH is larger than NL because vowels such as /i/, /I/, /ɛ/, /æ/ have the high second, third and fourth formant. Such frames are decided by the existence of the first formant." An interval *intv* is the vector of LSP's intervals, where $intv(i) = |p_{i+1} - p_i|$. If min (intv) < $F_s/4$ or if it is false that a < b, then vowels are detected because the LSP interval is narrow. (Page 999: Figure 4) (Note: "a < b" is the same as "a/b < 1", where "1" is "a prescribed threshold".) Thus, *Lee et al.* does disclose judging whether vowels are present in a speech signal when an interval on a frequency axis between LSP coefficients is less than a prescribed threshold.

Thirdly, Applicants cite Page 174, Left Column, of *Gersho et al.* as suggesting that that the bit rate is *necessarily fixed* as required for TDMA. Thus, Applicants argue that *Gersho et al.* does not specifically teach a variable rate encoder. Applicants posit

that *Gersho et al.* only theorizes about what *might be* possible without providing specifics. (*Applicants' emphasis*)

However, it is respectfully maintained that Applicants' interpretation of *Gersho et al.* is incorrect. Applicants' citation, indeed, notes that *Gersho et al.* goes on to say, "Nevertheless coders based on phonetic segmentation are well-suited for variable-rate coding." At best, Applicants' citation suggests only that TDMA without phonetic segmentation may not support variable rate coding. Still, *Gersho et al.* expressly suggests variable rate coding with phonetic segmentation. Generally, *Gersho et al.* notes it may not be cost-effective to exploit a voice activity factor in FDMA or TDMA, but variable rate coding is possible for CDMA and PRMA. Further, *Gersho et al.* states "TDMA can also be designed to benefit from voice activity patterns." (Abstract) Thus, *Gersho et al.* cannot be read to exclude variable rate coding, as implied by Applicants.

Finally, with respect to *Kang et al.*, Applicants argue that one of ordinary skill in the art would not have been realistically motivated to modify the vocoder of *Lee et al.* to include vowel templates because no benefit would have been gained by doing so. Thus, Applicants conclude, a *prima facie* case of obviousness is not established. This position is traversed.

Kang et al. suggests that template representations of LSP's provide an efficient data rate for voice and data encoding. (Column 1, Lines 49 to 58; Column 7, Lines 41 to 50) Template coding of vowels and consonants from LSP coefficients for voiced and unvoiced speech is equivalent to a codebook. Providing on the order of 100,000 templates for representing voiced and unvoiced speech produces refinement in

identification of voiced and unvoiced speech segments for more accurate encoding. Those skilled in the art know that each template is represented by an index specified by a number of bits. The total bit rate for speech coding is 800 bits per second in *Kang et al.* Thus, *Kang et al.* suggests it is effective and efficient to use templates for encoding speech and data.

Therefore, the rejections of claims 2, 3, 6, 7, 10, and 11 under 35 U.S.C. 103(a) as being unpatentable over *Lee et al.* in view of *Gersho et al.*; of claims 4, 8, and 12 under 35 U.S.C. 103(a) as being unpatentable over *Lee et al.* in view of *Gersho et al.*, and further in view of *Kang et al.*; and of claims 13 to 15 under 35 U.S.C. 103(a) as being unpatentable over *Lee et al.* in view of *Gersho et al.*, and further in view of *Das*, are proper.

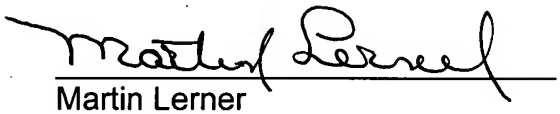
Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (703) 308-9064. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703) 305-9645. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 872-9314 for regular communications and (703) 872-9315 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-4700.

ML
January 26, 2005



Martin Lerner
Examiner
Group Art Unit 2654